



Program Evaluation - Learning how to determine whether your educational course, clerkship, or residency training program is “successful”

Steven J. Durning, MD

Paul A. Hemmer, MD, MPH

USU Department of Medicine



Acknowledgements

- **Louis N. Pangaro, MD**
- **Dario Torre, MD, MPH, PhD**



Program Evaluation: Overview

- Definitions
- Purposes
- Cardinal Questions
 - Why, What, Who, When, Where
 - **HOW: Framework**
- New or Existing Programs
- Examples, Activities throughout



Educational Goals



Curriculum



Evaluation

Individual **Program**

A horizontal line of text with a solid black arrow pointing downwards from the word 'Individual' to the word 'Program'.

Feedback / Grading



Phases in evaluation: Distinctions

- **Assessment** = making the observation
 - “assay” (measurement)
- **Evaluation** = determining value
 - in context of expected learner/program level
 - an educational “diagnosis”
- **Grading/Feedback** = taking an action
 - administrative/societal
 - educational intervention



Program Evaluation

- Systematic collection and analysis of information related to the design, implementation, and outcomes of a program, for the purpose of monitoring and improving the quality and effectiveness of the program.*
- System level of examination
- More than the sum of trainee data
- Defining “success” and what explains it

ACGME Glossary of Terms

Accreditation Council for Graduate Medical Education



Program Evaluation: Our Discussion Applies To:

- Entire educational program
 - Medical School, Residency
- A component
 - Basic Science yrs, Clinical yrs
 - Course, clerkship, year of residency
- An aspect/session
 - Lecture series, Direct observation
 - Instituting “home visits”



Program Evaluation: Why

- Accreditation bodies require it¹
- Determine strengths, weaknesses²
- Competing goals of stakeholders
- Obtain and/or justify resources

¹www.lcme.org

²Hayden, 2001



Program Evaluation: Goals Defining “Success”

- Educational programs deal with *change*
 - Program Evaluation should examine for change
- It is about *understanding* and *action*
 - To find the sources of strengths, weaknesses of the program, accreditation (External audience)
 - For decisions, local goals (Internal)
 - Trainees, Faculty, Program, System all contribute
- Relating “inputs” to “outputs”
 - And back again



Examples of Goal Questions

- “I would be happy if I knew: is there consistency of evaluation across medical student clerkship training sites?”
- “I would like to know: has reducing resident work hours affected student education?”



Individual Activity: 5 minutes

Exercise in Program Evaluation

Worksheet

Step 1

- Write down 1 Goal for your program
- Complete: “I would be happy about my program if I knew....”
 - Or: “I want/need to know.....”
- Be as specific as you can be:
 - Could you determine a Yes/No answer?



Large Group Examples of Goals from You



Activity: 5 minutes

- Do you want to revise your Goal Question?
- Do you want to pose a different question?



How can we help you get a quality answer to your question?

- Useful framework
 - Before, During, After (**How**)
 - Decide on the methods needed (**What**)
- You should be able know what you can attribute to the program, and what is simply due to the trainee
 - Is the program “successful” because of you or in spite of you?



Program Evaluation: How “Before, During, After” Framework

Before	During	After
<ul style="list-style-type: none">• Baseline characteristics• What the trainee brings to the program• What has been their experience	<p>Your program</p> <p></p>	<p>Outcomes</p> <p>This is also Table 1 in a Clinical Research study</p>



Program Evaluation: How Using the Framework

- Assessments in the framework
 - Main focus today
- Analysis:
 - Beyond scope of today's talk
 - Understanding of analytic methods informs assessments you'll choose
 - Quantitative or Qualitative analysis



Program Evaluation: What

- Multiple measures
 - No single measure can encompass all
 - Minimum of two assessments per “phase”
 - Quantifiable and Qualitative
 - Process and Product
- Prioritize: Essential or Desirable?
- Collect data under “controlled” conditions
- What are you already doing?
 - If collected, done well?



Program Evaluation: What Assessments in the Framework

- Process measures
 - Curriculum content (What we do with trainees)
 - Process measures help explain outcomes
 - E.g., if you create an educational intervention, you need to know how many received it
 - Process measures are typically part of the “During” phase
- Product measures
 - Outcomes (both primary and secondary)



Program Evaluation: What Assessments in the Framework

- Quantifiable
 - Numbers, counts (e.g., exams)
- Qualitative
 - Words (interviews, written comments)
 - Help explain variance in outcomes
- Both are key to
 - Finding unexpected results
 - Early warning of problems



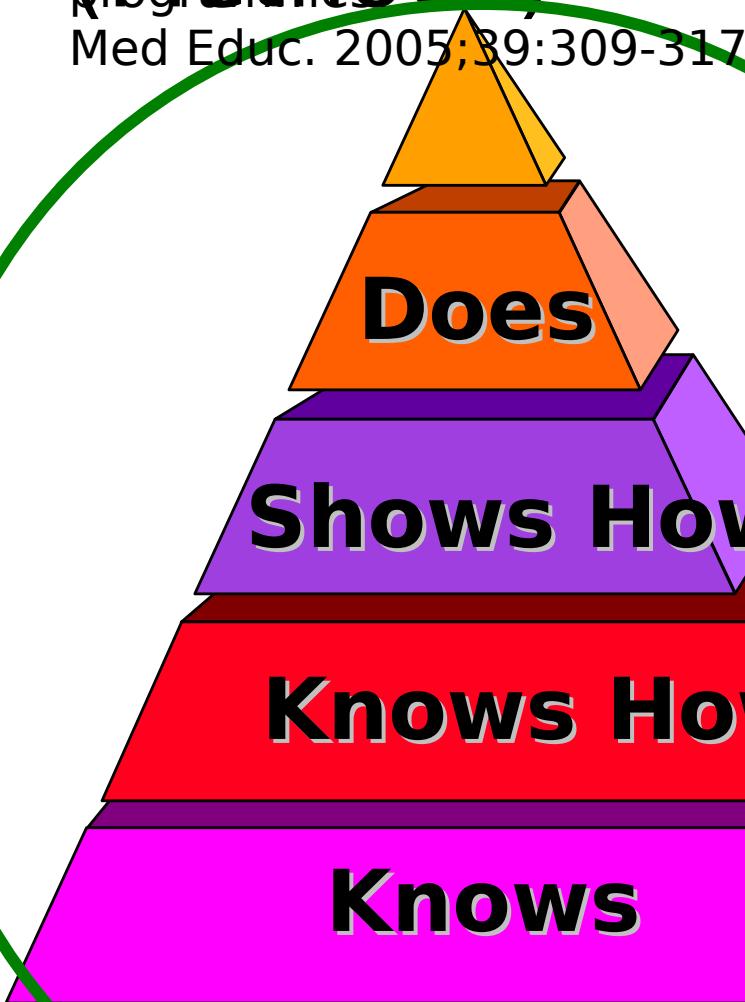
Program Evaluation: What Assessments in the Framework

- Reliable
 - Internal consistency, Signal to noise
 - Reproducible
- Valid
 - Confidence in the inferences drawn from information
 - Association with “outcome”
 - What are the sources of evidence to support validity?



Authenticity of Assessment

Assessing Professional Competence: from methods to programmes
Med Educ. 2005;39:309-317.



Descriptive Evaluation

(Teachers' words, workplace
Unannounced SPs, Chart
review)

Simulations, OSCEs,
mCEX

Simulations,
Simulators,
Prosthetics

MCQs/USMLE
vignettes

**Miller (Acad Med.
1990)**



Authenticity of Assessments

“Defense Against the Dark Arts”

- “It is not, however, always possible to test mastery of important skills in a standard examination format. Wizards, for instance, must be competent in using the Patronus charm to repel Dementors... Professor Lupin recruits a Boggart to simulate a Dementor when teaching Harry Potter how to conjure a Patronus charm.”
- “Professor Dolores Umbridge, an educational expert from the Ministry of Magic, makes the fundamental error of providing only a theory-based course for fifth year Defense Against the Dark Arts, a subject with performance-based assessment.”

Harry Potter and Assessment.

Jennifer Conn and Susan Elliott. The Clinical Teacher. 2005;2(1): 31-30



Before

During

After

What: Before (Baseline) Measures

- Identify, correct, control for differences
- What is available?
 - Do you have to create something?
 - Does a clerkship director know preclinical performance measures in order to adjust for baseline student differences across sites?
 - Are academic issues shared among directors?



Before

During

After

What: During Measures

- “Right” combination often program specific
- Examples
 - ACGME “toolbox”, exams, monthly evaluation forms, surveys, peer evaluations, attendance, portfolios, reflections



Before

During

After

What: After Measures Educational Outcomes: Kirkpatrick

- Level 1: learners' views/reaction
- Level 2: changes in what learners know
- Level 3: change in what learners can do
- Level 4: impact, benefit to clients or organization (Patients, stakeholders)

Kirkpatrick, D. L. (1967)



Before

During

After

What: After Measures--

Examples

- Board Certification/Examinations
- Placing graduates (fellowships, practice)
- Surveys (i.e. graduate or employer)
- Research and teaching productivity
- Academic appointments/advancement
- Disciplinary (i.e. NPDB, State Board)
- Patient Care (i.e. LDL, A1C targets, safety)



Before

During

After

Relate Before and During to After

- Do not get lost in focusing on Outcomes
- There should be a through-line connecting your After Measures to During and Before
- And:
 - Program evaluation may not be *linear*
 - Program evaluation does not end with “After” but should feed information back at all stages
 - “The only program that is not changing is a dead program.”



Program Evaluation: Individual Activity

Exercise in Program Evaluation

- **“What to evaluate” worksheets**
 - Steps 2-4
 - After, During, Before Measures
- **20 Minutes: Individual work**
 - Build on your Goal Question
- **Facilitators available**



Program Evaluation Small Group Activity

- **Exercise in Program Evaluation:**
“What to evaluate” worksheet
 - Steps 2-4
 - After, During, Before Measures
- **15 minutes small group debrief**



Program Evaluation: Who

- Gathering data can be performed by ALL
- Academic managers play primary role in interpreting data
- Consider all your stakeholders
 - Can expand resources, data for analysis



Program Evaluation: Who Stakeholders for Residency

- **Residents**
 - Education committee
- **Students**
 - Program critiques
- **Faculty**
 - Program resources, limitations
 - Overall resident strengths and weaknesses
- **Patients**
 - Real or SP
- **Peers, Nurses, etc.**
- **Chief Resident**
 - Resident problems/progress
- **Clerkship director**
 - Site resources/limitations
 - Faculty and resident morale
 - Faculty development
- **Oversight committees**
 - Academic actions



Program Evaluation: Who

- Don't go it alone
 - Create working groups, committees
 - Use variety of vantage points, different skills
 - Seek help
 - Statistical analyses
 - Qualitative analyses
 - Work within and across programs
 - Having others expect something is motivating!



Program Evaluation: Where Resources-What you have, What you need

- Time
 - Evaluation/feedback [verbal and written, interactive]
 - Faculty development (e.g. Frame of Reference)
 - Review/analyze/disseminate collected information
 - Complete annual evaluations/reports
 - To make scholarly work “scholarship”
- Human resources
 - Coordinator, clerical support, committees, faculty, ancillary
- Funding
 - Supplies, informatics, research, fac dev, offices, staffing



Program Evaluation: When

- Depends on program, resources
 - At least annually
 - Quarterly for recurring courses/clerkships
 - More frequently if implementing changes?
- Ideally: collect, analyze, interpret continually
 - Criteria helpful to “triage”, monitor information
 - Reminder of “what” data to collect, need mechanism for detecting problems
- **When** you perform PEv is **key** to finding concerns
 - **Yellow Flags:** closely follow
 - **Red Flags:** immediate attention



Examples of Flags

YELLOW

Test scores down 1-2 SD

Short program accreditation

Long resident hours

RED

Quantifiable

Test scores down 2 SD

Minimal accreditation/warning

Excessive resident hours

Qualitative

Learners not treated collegially

Report of learner abuse

Teachers give little feedback

Teachers ignore learners

Multiple “near” misses

Multiple sentinel events



Warning Flags: Example

- Geriatrics Home Visit
 - Internal medicine clerkship students visit patient known to them in the home
 - Geriatrics teaching modules provided
 - Students must have plan for visit
 - Students submit reflective write-up
 - **Required:** if not done, grade affected
- See Handout



ACTIVITY

Exercise in Program Evaluation

1) Steps 5-8 (15 min)

Handouts:

- 1) Resources for Program Evaluation
- 2) Red and Yellow Flags



Discussion (Small Groups)



Take Home

- Begin with defining SUCCESS
 - “I would feel really good about my program if I knew that...”
- Before, During, After Framework
 - Try for 2 measurements for each phase
 - Advocate collecting at least 3
 - Include qualitative and quantitative
 - Expect the unexpected (red, yellow flags)



Take Home

- Practicalities
 - Define needed resources (can you do it?)
 - Timing (include flags)
- Analysis
 - Define unit of analysis for consistency (site, yr)
 - Try to explain the outcome
 - ANOVA and linear or logistic regression
 - Qualitative



Getting started

- Start with existing programs
 - You HAVE to do this
- Don't go it alone—create group process
 - Departmental, Interdepartmental, School
 - Decide on what needs to be assessed
- Set deadlines, expectations
 - Annual reviews of courses, clerkships

